iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL
FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** V **Month of publication:** May 2025

DOI: https://doi.org/10.22214/ijraset.2025.68611

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

Birds Flu Disease Detection Using Machine Learning

Chanchal Inchulkar, Apoorva Thakre, Avantika Jengthe, Vaishnavi Kande, Prof. Gaurav Agrawal
Department of Computer Engineering, Cummins College of Engineering for Women, Nagpur, India

Abstract: Avian influenza, commonly known as bird flu, is a highly transmissible virus that predominantly impacts bird species, but in specific instances, it can also spread to humans. Identifying bird flu at an early stage is essential for limiting its transmission

among poultry and minimizing the chances of it affecting human populations. Symptom Recognition in Birds: Detection typically starts with observing clinical signs like coughing, sneezing, swelling, and reduced egg production in birds. These noticeable symptoms trigger further examination.. Molecular and Genetic Testing: PCR (Polymerase Chain Reaction): One widely used method for bird flu detection is RTPCR, which identifies specific genetic sequences from the virus.. Rapid Diagnostic Tests (RDTs): These tests allow for faster detection by identifying viral antigens in bird samples, offering quick initial results. Serological Testing: ELISA (EnzymeLinked Immunosorbent Assay): This technique detects antibodies produced in response to the bird flu virus, helping identify past or ongoing infections in bird populations.

Imaging and Data Interpretation: Cuttingedge AI techniques combined with image analysis can help evaluate patterns and irregularities in bird flocks that may signify infection. This is often coupled with data from sensors in automated farming systems to monitor bird health continuously.

Early Detection in Humans: For human detection, symptoms like fever, respiratory distress, and exposure to infected birds raise suspicion, followed by similar PCRbased lab tests to confirm infection. Surveillance Systems: National and global surveillance systems track outbreaks using data from wildlife monitoring, agricultural sites, and health agencies to detect and respond to bird flu outbreaks quickly. DataDriven

Models: AI and machine learning models play a growing role in predictive analysis, helping forecast outbreaks based on environmental, migratory, and climatic factors, offering a proactive approach to disease detection and control. Detection involves a multistep, integrated approach utilizing molecular diagnostics, serological assays, technologydriven surveillance, and predictive analytics to mitigate the spread of the virus effectively.

I. INTRODUCTION

The 1997 outbreak in Hong Kong SAR highlighted the pandemic risk of the H5N1 strain of Avian Influenza, offering new perspectives on how novel pandemic viruses could arise. Before 1997, pigs were thought to be the primary hosts for viral gene exchange, as their respiratory cells have receptors that accommodate both avian and human influenza viruses. However, the Hong Kong event demonstrated that humans, too, can contract the avian influenza virus, suggesting they could also act as hosts for the genetic exchange of the virus. Through these findings, H5N1 was then seen to have pandemic potential. Efforts have been made by researchers to control and eradicate the HPAI virus, yet the recurring losses in poultry remain a persistent threat to human health. Various factors contributing to the disease's spread can be analyzed to predict the onset and impact of future avian influenza outbreaks. Research into environmental factors, including land use, agricultural practices, trade activities, poultry populations, and farming methods, has identified these as key contributors to Tree, ANN, SMO Logistic Regression, kNN, LogitBoost. Integration with public health infrastructure and addressing computational efficiency for largescale implementations are also potential gaps. [7] Disease Prediction Using Machine Learning : The SVM algorithm classifies patient data by finding an optimal hyperplane that separates diseasepositive and diseasenegative cases. Enhancing model interpretability and integrating with healthcare systems for practical application are areas for further exploration. [8] Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks:ARIMA algorithm is used to model and predict outbreaks based on historical timeseries data, focuses on autocorrelations within the time series to predict future outbreaks.

Ability to handle nonlinear and complex patterns, the need for more diverse data, realtime outbreak prediction and model adaptability to changing conditions need further investigation.

[9] A global model of avian influenza prediction in wild birds: the importance of northern regions:The Random Forest algorithm aggregates the results from all trees for accurate predictions. Need for more comprehensive data from diverse regions, especially southern areas, to enhance model accuracy, exploring the integration of other algorithms and mproving he the introduction and spread of the disease.

II. LITERATURE REVIEW

[1] COVID19 Forecasting through Machine Learning: The ARIMA model has proven effective in predicting COVID19 progression, aiding in pandemic management and strategic response planning. Integrating diverse data sources, improving uncertainty handling, enabling realtime adaptation, focusing on longterm trends, and addressing ethical issues. Addressing these gaps will enhance the accuracy and practical utility of predictive models.

[2] A Framework for Avian Influenza Risk Prediction: The rule discovery algorithm, as demonstrated in an Indonesian case study, plays a vital role in identifying key factors and enhancing utbreak prediction and management capabilities.

Research gaps in avian influenza risk prediction include limitations in data comprehensiveness, the need for advanced modeling techniques, emerging risk factors, challenges in implementation, and the importance of interdisciplinary approaches.

[3] AIbased System for Swine Flu Prediction Using Naive Bayes: The Naive Bayes classifier uses historical data patterns to make probabilistic predictions, offering an efficient tool for the early detection and effective management of swine flu outbreaks. Include the need for more comprehensive datasets, particularly incorporating realtime and diverse environmental factors.

[4] Comparative evaluation of time models for predicting influenza outbreaks: application of influenza like illness data from sentinel sites of healthcare centers in Iran: SVM which classifies influenza outbreak data through optimal hyperplanes and ANN which models outbreak patterns using interconnected neurons to capture nonlinear relationships. Need for larger, more diverse datasets, improving realtime prediction accuracy, integrating these models into public health systems for practical application could be further investigated.

[5] Machine Learning for Dengue Outbreak Prediction: A Performance Evaluation of Different Prominent Classifiers:Naive Bayes, Decision Tree, ANN, SMO Logistic Regression, kNN, Logit Boost. Need for more comprehensive datasets that capture iverse environmental and social factors, improving odel accuracy in realworld scenarios, and by ntegrating predictive models with public health decisionmaking systems for proactive outbreak management.

[6] Predicting Malarial Outbreak using kNN Machine Learning and Deep Approach: A Review and Analysis: Naive Bayes, Decision i interpretability of Random Forest predictions are key areas for future research.

[10] Ensemble Approach for Zoonotic Disease Prediction Using Machine Learning Techniques: Gradient Boosting Machine + Random Forest Random Forest + Support Vector Algorithm. Need for more diverse datasets, and addressing the complexity of combining multiple algorithms, exploring model interpretability and scalability for practical, largescale deployment remains an area for further investigation.

[11] Identification of Disease Prediction Based on Symptoms Using Machine Learning: Random Forest, SVM, Decision Tree ,KNN, Naive Bayes. he limited exploration of rare and complex diseases, reliance on homogeneous symptom datasets, and the need for models that account for demographic and cultural variations.

[12] Towards Using Recurrent Neural Networks for Predicting Influenza like Illness: Case Study of Covid19 in Morocco : Long short term memory (LSTMs) are beneficial due to their ability to capture temporal dependencies in time series data. An area within a field of study that has not been explored or adequately addressed. It highlights a lack of existing knowledge or unanswered questions, indicating opportunities for further investigation.

III. ANALYSIS

PAPER NAME

Forecast and prediction of COVID19 using machine learning A framework for the risk prediction avian influenza occurrence: An Indonesian case study An Artificial Intelligent based System Naive Bayes for Efficient Swine Flu Prediction using Naive Bayesian Classifier

IV. METHEDOLOGY ACCURACY/ FINDING

Random Forest 91.63% Rule discovery algorithm The effect of each rule and its risk were computed For new set of symptoms the patient's diagnosis was done, and the prediction

of whether the patient has the disease or not was made 89.2% 88.9% 86.2% Comparative evaluation of time series models for predicting influenza outbreaks: application of influenza like illness data from sentinel sites of healthcare centers in Iran Machine Learning for Dengue Outbreak Prediction: A Performance Evaluation of Different Prominent Classifiers
SVM
ANN
Random Forest
Naive Bayes
Decision Tree
ANN
SMO
Logistic Regression
KNN
Logit Boost
KNN
SVM
ANN
Naive Bayes
Random Forest
XGboost
Logistic Regression
Support Vector Machine Heart Dataset75.4%,
Diabetes Dataset78.6%
Liver Dataset75.9%
Random Forest model's Mean Square
Error (MSE), was less than both th
ARIMA's simulated retrospective MS
and the prospective ARIMA's MSE
ARIMA (1,0,1): AIC159.16 Error%=
29.7% 81% 84% 81% 80% 85% 75% 92 86.21% 92.69% 25.83% 91.69% 93.94% 96.26% 92.44%
Predicting Malarial Outbreak using Machine Learning and Deep
Approach: A Review and Analysis
Disease Prediction Using Machine Learning
Comparison of ARIMA and Random Forest time series models for prediction ARIMA of avian influenza H5N1 outbreaks Random Forest, Modeling Roles and of Factors in Outbreaks of Highly Pathogenic Avian Influenza H5N1 Towards Using Recurrent NeuralNetworks for Predicting Influenza Memory like Illness: Case Study of Covid19 in Morocco Identification of Disease Prediction
ARIMA, SARIMA
Epidemic prediction and study to make predictions on Covid19 pandemic growth in Morocco, using LSTM model was done Random Forest, SVM, 94.12%
Long short term Based on Symptoms Using Machine Learning
Decision Tree
KNN
Naive Bayes
95.01% 95.13% 92.68% 94.25% 92.90% 93.85% 97.64% 44.6% 67.2%
Disease Prediction using Machine Learning
Naive Bayes
Decision Tree
Random Forest
Gradient Disease
Boosting Machine +

Random
Forest
Random forest
Ensemble Approach for Zoonotic
Disease Prediction Using Learning Techniques
A global model of avian influenza prediction in wild birds: the importance of northern regions 79%

V. CHALLENGES

1) Data Limitations: Insufficient availability of well labeled and highquality datasets.
2) Symptom Variability: Variations in symptoms observed across different species.
3) High Dimensionality: Managing complex, multi feature data.
4) Feature Extraction: Identifying relevant biological and environmental features.
5) Privacy Concerns: Issues in using sensitive biological data.
6) Environmental Factors: Difficulty modeling ecological influences.
7) Demanding systems.
8) Model Interpretability: Making sure that the models are easy to understand and explain.
9) CrossSpecies Transmission: Complicating detection models.
10) Virus Evolution: Constant need to update models due to mutations on backyard poultry like ducks, geese, chickens, etc.

VI. FINDINGS AND RESULTS

A future outbreak of HPAI can be predicted using the outlined methodology. Figure 1 illustrates the steps required to accomplish this goal. Initially,data collection would take place, followed by preprocessing to ensure that no critical features are missed. Feature extraction would include scaling and translating the data as necessary. Rows with missing data may be discarded to maintain dataset integrity. The dataset would be split randomly into training and testing sets. The training data would be used to train the model, while the test data would be reserved for evaluating the model's performance.

To prevent overfitting, the test data would be utilized only once. Additionally, a validation dataset may be created to allow for repeated evaluations of the model. These two approaches aim to enhance computational efficiency, minimize overfitting, and eliminate noise from the data.

At this stage, two datasets will be generated: one containing all attributes and another with a reduced attribute set. The reduced dataset would also be randomly split into training and testing sets. The datasets obtained thereafter would be applied to different classifiers, in this case separately to simple and then ensemble classifiers. Model performance would be evaluated using both the full set of attributes and the reduced attribute set.

A confusion matrix would be used to calculate the accuracy, sensitivity, specificity, recall, and precision for each model. The model demonstrating the highest accuracy would be selected for predicting avian influenza outbreaks.

RealTime Detection: Computationally pandemic.

The dataset used would include maximum features that impact the occurrence of Avian Influenza, such as meteorological factors and duck and chicken populations. In previously done research, chickens and pigs, as important transmitter of avian influenza, have been studied widely. So, for this research, duck and geese populations will also be considered for the prediction of avian influenza and its spread in the Asian region. Results using classifiers other than the ones used in earlier studies will also be explored to predict avian influenza disease. Several different climatic variables will also be considered as attributes to study the refinement of the predictability of the proposed model and also to verify if the impact of the climatic variables is in accordance or in contradiction to prior research.

VII. CONCLUSION

The rapid transmission of the disease could potentially lead to an increase in mutation events and genetic reassortment, making it harder to control. This raises the likelihood of avian influenza developing into a pandemic in the future. For the prevention of the devastation caused by the disease, forecasting the disease could be a step in the right direction.

Accurately forecasting potential outbreaks is critical for effective disease management and control, helping to minimize the devastating consequences, both in terms of lives lost and economic impact.

Avian Flu may present a potential challenge by hampering he development of a particular region and possibly the world, much similar to the current COVID19 time series models for prediction of avian influenza H5N1 outbreaks:

Comparison of ARIMA and Random Forest PROCESS WORK

1) HOME PAGE User Response Page
2) DATA/RESPONSE Gathering data from the farm
3) DATA PREPROCESSING Training the machine learning model
4) FEATURE TO CLOUD RealTime Disease detection
5) DISEASE DETECTION Run live image from the model
6) ALERT & REPORTING Reporting the disease if generated

REFERENCES

[1] Forecast and prediction of COVID19 using machine learning https://ieeexplore.ieee.org/document/9377038
[2] A framework for the risk prediction avian influenza occurrence: An Indonesian case study: https://www.nature.com/articles/s41598020758897
[3] An Artificial Intelligent based System for Efficient Swine Flu Prediction using Naive Bayesian Classifier: https://ijisae.org/index.php/IJISAE/article/view/4073
[4] Comparative evaluation of time models for predicting influenza outbreaks: application of influenza like illness data from sentinel sites of healthcare centers in Iran: https://bmcresnotes.biomedcentral.com/articles/10.11 86/s131040194393y
[5] Machine Learning for Dengue Outbreak Prediction: A Performance Evaluation of Different Prominent Classifiers: https://www.informatica.si/index.php/informatica/arti cle/view/1548/0
[6] Predicting Malarial Outbreak using kNN Machine Learning and Deep Approach: A Review and Analysis https://ieeexplore.ieee.org/document/8724266
[7] Disease Prediction Using Machine Learning:https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8896 926/ https://pubmed.ncbi.nlm.nih.gov/25123979/
[8] A global model of avian influenza prediction in wild birds: the importance of northern regions:https://pubmed.ncbi.nlm.nih.gov/23763792/
[9] Ensemble Approach for Zoonotic Disease Prediction Using Machine Learning Techniques: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC 9607465/
[10] Identification of Disease Prediction Based on Symptoms Using Machine Learning: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC 8896926/
[11] Towards Using Recurrent Neural Networks for Predicting Influenza like Illness: Case Study of Covid19 in Morocco https://www.sciencedirect.com/science/article/pi i/S0925231221015150
[12] Welekar, A.R. and Patif, M.E., Experimental Investigation of Segmentation Methods for Segmenting Fruit Images in NaturalEnvironment. In Recent Advances in Science, Engineering & Technology (pp. 355362). CRC Press.
[13] Amit R. Welekar, Manoj Eknath Patil. (2022). Proposing Improved FCM based Method for Segmenting Irregular Shaped Fruit Image Captured in Natural Light. Mathematical Statistician and Engineering Applications, 71(4), 9804–9813. https://doi.org/10.17762/msea.v71i4.1784